# SHORT TERM WIND POWER FORECASTING USING ADAPTIVE NEURAL FUZZY INFERENCE SYSTEMS

Peter Johnson, Michael Negnevitsky and Kashem M. Muttaqi

Centre for Renewable Energy and Power Systems
University of Tasmania,
Hobart, TAS, 7001, Australia

## Abstract

As the global political will to address climate change gains momentum, the issues associated with integrating an increasing penetration of wind power into power systems need to be addressed. This paper summarises the current trends in wind power and its acceptance into electricity markets. The need for accurate short term wind power forecasting is highlighted with particular reference to the 5 minute dispatch interval for the proposed Australian Wind Energy Forecasting System. Results from a case study show that Adaptive Neural Fuzzy Inference System (ANFIS) models can be a useful tool for short term wind power forecasting providing a performance improvement over the industry standard "persistence" approach.

## 1. Introduction

In addition to the various ecological and sustainability arguments, the recently published Stern Review on the economics of climate change [1] provides clear economic justification for the costs of action to minimise the worst impacts of climate change. Power generation accounted for 24% of greenhouse gas emissions in the year 2000 and the Stern Review highlights the fact that support policies for low carbon power generation technologies will be critical. Wind power is a maturing renewable energy technology with various multi-megawatt turbines now commercially available for installation in utility scale wind farm configurations. Wind power thus represents a commercially viable, renewable energy resource which should be considered in any future power generation mix.

The Global installed wind power capacity continues to grow at around 25% per year with more than 74GW installed world wide at the end of 2006 [2]. The current global installed capacity is expected to double by the end of this decade with the European Union leading the way followed by North America and Asia.

Australia has great potential for wind farm development with some of the world's best wind resources and highest capacity factors. At the end of 2006, there was 817MW of installed and operating wind farm capacity with an additional 521MW of wind farm projects under construction or nearing completion and a further 2100MW of projects with planning approval [3]. In the absence of an extension of the Federal Government's Mandatory Renewable Energy Target (MRET) scheme, various state governments have announced their own schemes to provide political support to the wind energy industry. Wind farm development in these states is responding according to the incentives provided.

As wind power continues to develop as a significant renewable power generation resource, a number of problems arise with integration of that resource into existing power systems and electricity markets. In this paper, the need for accurate short term forecasting of wind power generation is highlighted. A proposed approach using a hybrid intelligent system for wind power forecasting is discussed and results from an initial case study are presented and discussed.

### 1.1 Wind Power and Electricity Markets

Due to the inherent variability of wind, the power generation from wind farms is neither constant nor schedulable. As the penetration of wind generation increases, this unpredictability and variability has both technical and commercial implications for the efficient planning and operation of power systems [4].

Around the world, non-schedulable power generation such as that produced by wind farms is accepted into electricity markets in a variety of ways. In some cases, government regulated feed-in pricing is set for variable, renewable sources while in other markets, all participants must compete on equal terms.

In Australia, the wholesale electricity market operator responsible for system security and operation of the deregulated National Electricity Market (NEM) is the National Electricity Market Management Company (NEMMCO). In this market, a large proportion of power is hedged via long term financial contracts between generators and retailers. NEMMCO sets the wholesale pool price for each five minute dispatch interval based on generator bidding and settlement takes place on a thirty minute trading interval basis [5].

Wind power is currently accepted straight into the Australian NEM with wind power generators able to receive the market pool price. Additional revenue can be raised by wind farm operators under the MRET scheme through the sale of Renewable Energy Certificates (RECs).

For the purposes of ensuring system security and dispatching generation capacity to match demand, NEMMCO treats wind power as a stochastic 'negative demand' resulting in a modified peak demand profile. This links wind power forecasting to demand forecasting and the importance of accurate short term wind power forecasting becomes more significant as the penetration of wind power increases.

NEMMCO are currently managing the development of an Australian Wind Energy Forecasting System (AWEFS) to meet a set of published functional requirements [6]. These functional requirements highlight the importance of accurate short term wind power forecasting for the 5 minute dispatch interval and also call up forecast requirements for 11 subsequent 5 minute pre dispatch intervals out to 60 minutes ahead.

This paper focuses on the problem of predicting wind farm power generation output in the 'near real time frame' of 5 minutes ahead corresponding to NEMMCO's 5 minute dispatch interval.

### 1.2 Wind Power Forecasting

Numeric Weather Prediction (NWP) methods are well established for wind forecasting with a prediction horizon of several hours or more. The corresponding forecast wind speeds can be converted to an approximate power output from the wind farm of interest. NWP models can also be specifically tuned using accurate Digital Elevation Models (DEMs) and Model Output Statistics (MOS) corrections for shorter time frames but they have proven to be unsuitable for the very short term or 'near real time frame'.

For very short term time frames, available forecasting techniques range from the "industry benchmark" persistence approach to statistical methods and methods based on the use of Artificial Neural Networks (ANNs).

The persistence model relies on the high correlation between the current and forecast values in the short term and simply equates the forecast value to the current observation (i.e. predicts no change over the forecast interval). This approach becomes less valid as the forecast interval increases but persistence has proven to be a useful first approximation for short term wind and wind power forecasting and provides a benchmark against which to compare alternative techniques.

In the 'near real time frame', some work has been undertaken on the application of neural networks to model and estimate wind turbine output [7, 8]. However the specific tuning required and lack of general portability of these approaches can be restrictive.

To date, no single approach has gained widespread industry acceptance and the simple persistence model is often used even though more accurate predictions can be achieved.

## 2. Proposed Approach

In the field of intelligent systems, fuzzy logic and neural networks are natural complementary tools. These tools can be combined to produce a neuro-fuzzy system which is functionally equivalent to a fuzzy inference model. A neuro-fuzzy system can be trained to develop the IF-THEN fuzzy rules and adjust the Membership Functions for the associated inputs and outputs. One such system architecture, referred to as an Adaptive Neuro Fuzzy Inference System or ANFIS was originally proposed by J.-S. R. Jang. The six layer, feed forward, neural network structure of a simple two input, 2 Membership Function ANFIS is illustrated in Figure 1.

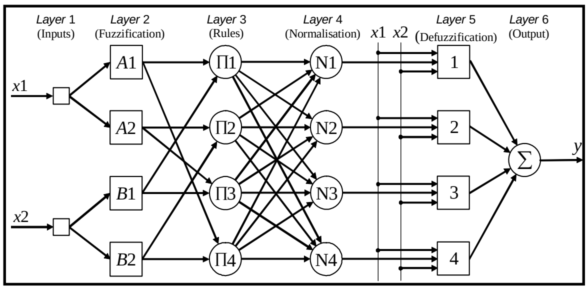


Figure 1: Example ANFIS Structure
(2 Inputs, 2 Membership Functions)

A number of variables must be defined when constructing an ANFIS model. These variables include; the type and number of inputs, the type (shape) and number of Membership Functions associated with each input (in the fuzzification layer), the size of the training set and the number of training epochs or iterations.

ANFIS models have been successfully applied to the prediction of chaotic time series and recent research has demonstrated that such models may have some merit for the prediction of wind vectors with a prediction horizon of 2.5 minutes ahead [9]. For this study, the application of an appropriate ANFIS configuration is proposed for predicting total wind farm power output in the 'near real time frame'.

## 2.1 Model Performance Evaluation

The need for a standardised approach to the performance evaluation of wind power prediction systems was recognised as part of the ANEMOS project in Europe [10] and appropriate forecast accuracy measures were also addressed in NEMMCO's functional requirements specification for the AWEFS [6]. The standard nomenclature applied to these measures is as follows ;

$P_{inst}$ = wind farm installed capacity

$k = 1,2..,k_{\max}$ - Prediction horizon (number of discrete time steps)

t = time origin (current time)

$N$ = number of data used for the model evaluation

$P(t+k)$ = measured (actual) power at time, t+k

$\hat{P}(t+k\,|\,t)$ = power forecast for time t+k made at time origin, t

$e(t+k\,|\,t)$ = error at time t+k for the prediction made at the time origin, t

Using the notation above, the prediction error for a particular prediction horizon, k is given by;

$$e(t+k\,|\,t) = P(t+k) - \hat{P}(t+k\,|\,t) \qquad (1)$$

The discrete time step associated with the variable k also needs to be defined in accordance with the time frame of interest. For the "near real time" prediction horizon of interest here a discrete time step of 1 minute is considered.

A number of error measures can be considered for the performance of a prediction model over the whole of an evaluation or test period. These include the average error (BIAS), the mean absolute error (MAE), the root mean squared error (RMSE) and the standard deviation of errors (SDE).

Each of these basic error measures can be normalized to the installed capacity of the wind farm under consideration. Such normalized error measures can be expressed as percentages and are used to provide an error measure which is independent of a particular wind farm size.

In line with the requirements for the AWEFS and published standard protocols [6, 10], a basic set of forecast performance measures were used to assess and compare models in this study. The two most important error measures examined were the Normalized Mean Absolute Error (NMAE) and the Normalized Root Mean Squared Error (NRMSE), which are defined as follows;

$$NMAE(k) = \frac{\sum_{t=1}^{N} |\,e(t+k\,|\,t)\,|}{N.P_{inst}} \qquad (2)$$

$$NRMSE(k) = \sqrt{\frac{\sum_{t=1}^{N} \left(e(t+k\,|\,t)\right)^2}{N}} \Big/ P_{inst} \qquad (3)$$

Performance comparisons can be expressed as percentage improvements compared to a reference model as follows;

$$imp\%_{ref,EC}(k) = \frac{EC_{ref}(k) - EC(k)}{EC_{ref}(k)}.100\% \qquad (4)$$

Where EC is the relevant Evaluation Criterion (eg, NMAE, NRMSE, etc.)

The "base line" reference model for improvement comparison associated with short term wind power prediction models is the simple persistence model as defined below;

$$\hat{P}_P(t+k\,|\,t) = P(t) \qquad (5)$$

### 2.1.2 Performance Comparison Methodology

The methodology used to test and compare the performance of wind power prediction models should involve the use of a test data set which is independent of any data used to train, tune or validate the model.

When looking to build and test a prediction model, two suitably diverse data sets, a training data set and a testing data set should be defined and the test data should be quarantined for performance testing only. The training data set (which may include validation data) should be used to build, train or tune the model while the testing data set (which is not used in any way for model development) should be used exclusively for measuring model performance based on the error measures considered.

When tuning and comparing the performance of competing models, the basic principles of good experimental technique need to be employed. This includes changing only one variable at a time and using the same testing data set and error measures to obtain valid comparisons.

## 3. A Case Study

In order to develop and test ANFIS based models to predict the output from a wind farm in the 'near real time frame', a suitable time series data set is required. This time series data set should be large enough to provide suitable training and testing data sets and these separate data sets should be sufficiently diverse to represent the range of output behaviour of the wind farm involved. In addition, the sampling rate of the time series data should be fast enough to ensure that output variations in the time scales of interest are captured.

The availability of wind farm production data with suitable fine time resolution has proven to be difficult to obtain from local operators, however a limited data set from a wind farm in the mid-west of the USA was available and was used for the case study presented here.

The wind farm associated with this case study comprised 150 turbines rated at 750kW each, providing a total installed capacity, $P_{inst}$=112.5MW.

### 3.1 Initial Data Analysis

The data set investigated comprised time stamped, 1 second samples of total wind farm power output over 14 contiguous days. The first 10 days of data was set aside for model training with the last 4 days quarantined for performance testing.

The data was initially plotted on a number of time scales and the plots were inspected to identify important features of the time series. Plots of the entire data set and a 30 minute "snapshot" of data are shown below in Figures 1 and 2.

By inspecting various 30 minute snapshots (similar to Figure 1) across the entire data set, a qualitative observation was made that data samples at intervals between 30 seconds and 1 minute appear to provide reasonable tracking of the output response. Periods of highly stochastic or "noisy" variations were also noted when the output power of the wind farm was at the extremes of the operating range.

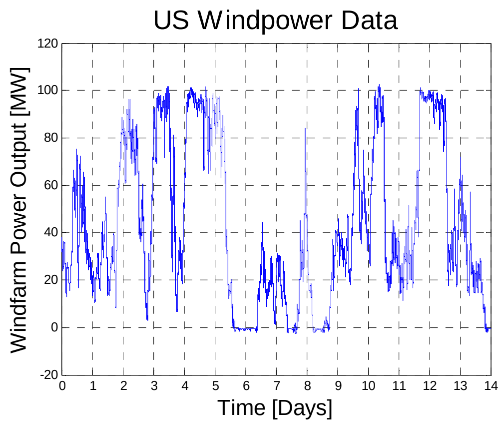


Figure 1: Complete Data Record

Preliminary spectral analysis of the time series was undertaken using Fast Fourier Transforms (FFTs) and plotting the associated power spectral density. This analysis indicated that the power output time series was generally over sampled with minimal power spectral density for frequencies higher than those corresponding to a 10 second sampling rate.

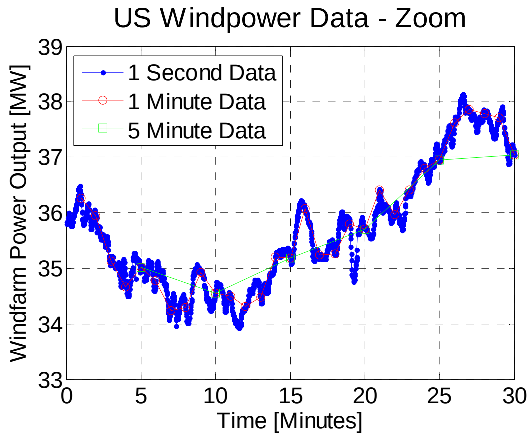


Figure 2: Thirty Minute Snapshot

### 3.2 Input Data Configuration For ANFIS Model

Initial ANFIS model trials concentrated on the prediction of wind farm power output 5 minutes ahead (in line with the first dispatch interval used in the Australian NEM). These initial investigations revealed that reasonable ANFIS model performance was not possible using the actual values of wind farm power output as inputs to predict a corresponding future output value. This was consistent with findings from previous research using ANFIS models to predict power demand and wind vector magnitude [9, 11]. This inability to adequately model the raw data may be attributed to the large range of output values involved and the inability of the ANFIS models trialled to be adequately trained over that range.

The use of difference values was then investigated. This involved using the difference between successive wind farm power output values as inputs to predict the corresponding future difference value. Differencing the time series data in this way is analogous to differencing to achieve stationarity for statistical models. Such differencing effectively reduces the range of possible values involved [9]. The predicted difference value is then converted to an absolute prediction value by adding it to the corresponding current value.

Initial trials using difference values in a variety of input configurations for an ANFIS model indicated that there was some merit in using 1 minute difference inputs to predict wind farm power output 5 minutes ahead.

### 3.3 Effect of Varying ANFIS Model Parameters

#### 3.3.1 Effect of the Number of Training Epochs.

Starting with a simple 3 input structure with 2 "bell" shaped membership functions per input, the effect of varying the number of training epochs was investigated. No significant improvement in the model performance was observed as the number of training epochs was increased in steps from 1 to 100 suggesting that suitable training was achieved with 1 epoch.

### 3.3.2 Effect of the Number of Inputs.

With all other parameters fixed and using the same training and testing data sets, the number of 1 minute difference inputs was varied to observe the effect on overall model performance and training time. In Figure 3 the percentage performance improvement of the ANFIS model over the persistence model for a 5 minute prediction horizon is plotted against the number of inputs to the ANFIS model. The performance improvement is measured using the NMAE and the NRMSE as error criteria, i.e. $imp\%_{persistence,NMAE}(5)$ and $imp\%_{persistence,NRMSE}(5)$. The corresponding ANFIS model training times are plotted on the same axes for reference.

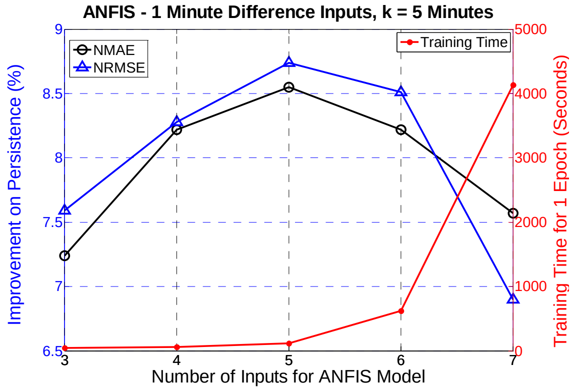


**Figure 3: Effect of Varying the Number of Inputs**

A clear peak in performance is observed with 5 inputs for this data set and model structure. As the number of inputs is increased beyond 5, not only does the performance decrease but the training time associated with the model increases exponentially.

### 3.3.3 Effect of the Input Interval.

Fixing the number of inputs at 5 and keeping all other parameters unchanged, the interval between inputs was then varied to investigate the effect on model performance. The results of this investigation are summarised in Figure 4 with performance improvement measured as before.

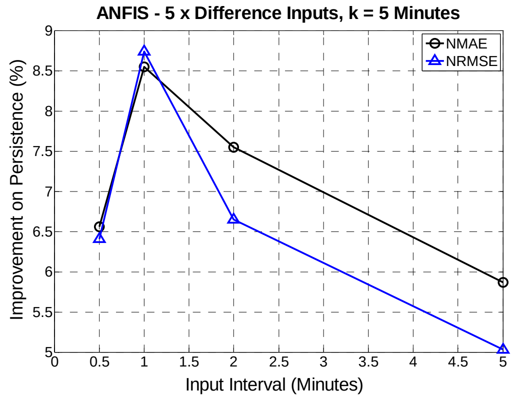


**Figure 4: Effect of Varying the Input Interval**

These results illustrate that the best input interval to use for this data set and model structure is in the order of 1 minute.

### 3.3.4 Effect of Membership Function Variations.

With the model structure fixed at 5 x 1 minute difference inputs, the effect of varying the type and number of the membership functions (MFs) associated with each input was then investigated.

Firstly, the number of MFs associated with each input was increased from 2 to 3. This resulted in a significant increase in training time from approximately 2 minutes to 2.5 hours due to the increased number of parameters in the model. The net effect of this increase was a slight **decrease** in the model performance using the same training and testing data sets.

Reverting to 2 MFs per input, the effect of changing the MF shape was investigated. The MFs types trialled comprised three smooth transition types (Bell, Gaussian and Pi) and two piecewise linear MFs (Triangular and Trapezoidal). In each case the difference in model performance was less than 0.01% indicating that MF shape is not a significant factor.

## 3.4 ANFIS Model Performance

From the results summarised above, the preferred ANFIS model structure for this data set and prediction horizon can be seen to comprise 5 x 1 minute difference inputs with two "bell" shaped Membership Functions per input. A comparison of the key performance measures of this preferred ANFIS model with the corresponding persistence model is provided in Table 1.

**Table 1: ANFIS Model Performance Summary**

| Error Criteria | ANFIS | Persistence | % Imp |
|---|---|---|---|
| NMAE | 1.39% | 1.52% | 8.55% |
| NRMSE | 1.98% | 2.17% | 8.75% |

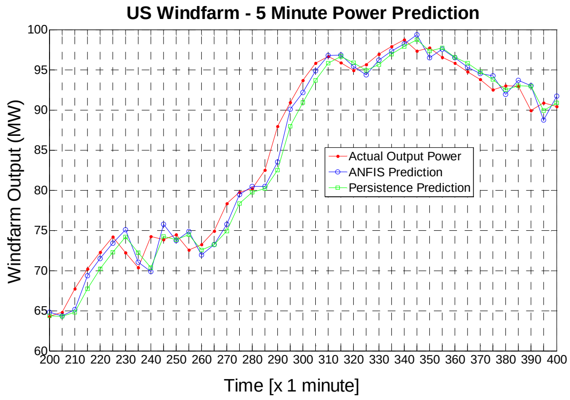


**Figure 5: Prediction Performance Snapshot**

A sample snapshot plot of predicted values compared to the actual power output times series values is shown in Figure 5.

## 4. Discussion and Conclusions

For the test case considered here, an ANFIS model configuration provided an overall performance improvement of more than 8% over persistence for 5 minute ahead predictions of the total wind farm power output. The prediction performance snapshot shown in Figure 5 indicates that the ANFIS model tends to provide markedly better predictions than persistence during periods when the wind farm power output is steadily increasing or decreasing but the ANFIS predictions tend to overshoot slightly at local maxima or minima in the time series.

While there was a notable performance improvement of the ANFIS model over persistence, in absolute terms the error performance of the persistence approach for a 5 minute prediction horizon was remarkably good. However, any improvement in prediction accuracy is worth pursuing and the ANFIS modelling approach investigated in this study showed promise.

The relatively short length of the time series data set available for this case study is recognized as a potential limitation given that it may not provide sufficient diversity to suitably represent all wind farm production scenarios for model training and testing. The limitation in the time series length is offset to some degree by the inherent over sampling of the data This oversampling enabled a suite of data sets to be constructed at 1 second offsets. This increased the number of available training sets which could be presented (in a randomised order) to train the model. The portability of this approach still needs to be tested over further case studies and other suitably diverse, fine resolution wind farm output data sets are being sought.

Extension of this forecast modelling approach to longer prediction horizons (e.g. out to 30 minutes or beyond in steps of 5 minutes) will be the subject of further research. Some initial attempts were made to apply the same model structure (5 x 1 minute difference inputs with 2 MFs per input) to 10, 15 and 20 minute prediction horizons. The performance of these models demonstrated prediction performance marginally better than persistence in absolute terms but the percentage improvement decreased as the prediction horizon increased and the persistence approach became less accurate.

The application of time series spectral analysis, correlation analysis and appropriate pre-filtering are also areas for further research. These tools may aid in developing a universal methodology to select the best model type and associated model variables for a given prediction horizon. Such a general methodology may include a range of modelling approaches including statistical models, neural network models or hybrid intelligent systems such as the one investigated here.

In summary, for market driven power systems there is a need for accurate wind power forecasting in the short term or 'near real time frame'. The application of an ANFIS hybrid intelligent system has been shown to provide improvement over the industry standard persistence approach for a look ahead period of 5 minutes in the case study investigated. Potential extensions of this approach have been proposed and some future areas of research have been outlined.

## 5. References


[1] N. Stern. *Stern Review: The Economics of Climate Change*. 2006 [cited 10 November 2006]; Available from: http://www.hm-treasury.gov.uk/independent_reviews/stern_review_economics_climate_change/stern_review_report.cfm.

[2] GWEC. *Global Wind 2006 Report*. [cited 25 July 2007]; Available from: http://www.gwec.net/.

[3] Auswind. *House of Representatives Inquiry into developing Australia's non-fossil fuel energy industry - Australian Wind Energy Association Submission*. 2007 [cited 25 July 2007]; Available from: http://www.auswind.org/auswea/index.html.

[4] G. N. Bathurst, J. Weatherill, and G. Strbac, "Trading wind generation in short term energy markets," *IEEE Transactions on Power Systems*, vol. 17, pp. 782-789, 2002.

[5] M. Negnevitsky, P. Johnson, and S. Santoso, "Short term wind power forecasting using hybrid intelligent systems," in *IEEE/PES General Meeting (panel paper submitted for publication)*: IEEE, 2006, pp. 4.

[6] *Functional requirements for AWEFS and NEM forecasting processes*. National Electricity Market Management Company Limited.18 August 2006 [cited October 10 2006]; Version 1.0:[ROI support documentation]. Available from: http://www.nemmco.com.au/dispatchandpricing/awefs.htm.

[7] S. Kelouwani and K. Agbossou, "Nonlinear model identification of wind turbine with a neural network," *IEEE Transactions on Energy Conversion*, vol. 19, pp. 607-612, 2004.

[8] L. Shuhui, D. C. Wunsch, E. A. O'Hair, and M. G. Giesselmann, "Using neural networks to estimate wind turbine power generation," *IEEE Transactions on Energy Conversion*, vol. 16, pp. 276-282, 2001.

[9] C. W. Potter and M. Negnevitsky, "Very short-term wind forecasting for Tasmanian power generation," *IEEE Transactions on Power Systems*, vol. 21, pp. 965-972, 2006.

[10] H.Madsen, G.Kariniotakis, H. A. Nielsen, T. S. Nielsen, and P.Pinson, "A Protocol for Standardizing the Performance Evaluation of Short-Term Wind Power Prediction Models," Deliverable 2.3, 2004.

[11] C. Potter, "Development of a very short-term prediction system for wind power generation," *PhD Thesis in School of Engineering*, University of Tasmania, Hobart: 2006, pp. 177.